Forum for Linguistic Studies | Volume 07 | Issue 02 | February 2025

BILUNGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

The Effect of Using Real-Time Magnetic Resonance Imaging on the Perception of EFL Vowels

Faisal Aljasser

Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah 52571, Saudi Arabia

ABSTRACT

This study investigated the influence of real-time Magnetic Resonance Imaging (rtMRI) videos as an innovative tool for enhancing learners' perception of English vowel sounds among Arabic-speaking learners of English as a Foreign Language (EFL). Utilizing a pretest-intervention-posttest design, the study involved forty participants who were equally divided into two groups: an experimental group and a control group. The experimental group received targeted perception-based instruction that incorporated rtMRI videos, delivered in three weekly sessions of thirty minutes each. In contrast, the control group underwent a similar instructional approach but without the use of rtMRI technology. To assess learners' progress, two tests were administered during both the pretest and posttest phases. The first test focused on the subjects' ability to associate vowel sounds with their corresponding articulatory features, while the second test evaluated their overall perception of vowel sounds. Analysis of the post-test results revealed significant improvements in the vowel-feature association test exclusively for the experimental group. Implications of these findings are discussed.

Keywords: EFL Vowels; Perception; Pronunciation; rtMRI; Visual Cues

*CORRESPONDING AUTHOR:

Faisal Aljasser, Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah 52571, Saudi Arabia; Email: jasr@qu.edu.sa

ARTICLE INFO

Received: 6 December 2024 | Revised: 19 December 2024 | Accepted: 20 December 2024 | Published Online: 16 January 2025

DOI: https://doi.org/10.30564/fls.v7i2.8001

CITATION

Aljasser, F., 2025. The Effect of Using Real-Time Magnetic Resonance Imaging on the Perception of EFL Vowels. Forum for Linguistic Studies. 7(2): 1–15. DOI: https://doi.org/10.30564/fls.v7i2.8001

1. Introduction

Learning second language (L2) phonology is a complex process influenced by a variety of linguistic and cognitive factors. For Arabic-speaking learners of English as a Foreign Language (EFL), the challenge of accurately perceiving and producing English vowels is especially pronounced due to the differences in vowel systems between Arabic and English. English, with its rich inventory of vowel sounds, presents a significant learning hurdle for Arabic speakers, whose native language has a more limited vowel system. This study aims to investigate how rtMRI can be utilized to enhance the perception of English vowels among Arabic-speaking EFL learners. By providing visualizing articulatory movements, rtMRI offers a novel approach to understanding and improving vowel perception in L2 learners.

Perception of EFL vowel sounds requires identifying the properties that define them, with closeness, frontness, and rounding being key features. Closeness pertains to the degree of jaw openness, evident in pronouncing /ɑː/ as in *car* with wide jaws and an open mouth. Rounding involves observable lip movements, as seen in the contrast between *feel* and *fool*. However, the feature *frontness* refers to the position of tongue elevation on a horizontal axis. For example, the vowel in *feel* is considered a front vowel because the highest point of the tongue is towards the front of the mouth, while the vowel in *fool* is categorized as a back vowel because the highest point of the tongue is towards the back of the mouth [1]. While learners can visually observe the first two features, frontness cannot be visually observed. Therefore, this study aims to explore the impact of using rtMRI videos as a novel tool to enhance learners' perception of EFL vowel sounds. RtMRI offers real-time visualization of articulatory movements, providing learners with dynamic visual feedback that was previously unavailable, potentially increasing their awareness and improving their ability to accurately perceive and distinguish EFL vowels.

Traditionally, teaching speech sounds has relied heavily on production-based instruction using corrective feedback and visual aids like vowel charts. However, these methods may be limited in their effectiveness because similar-sounding vowels can be difficult for non-native listeners to perceive solely through vowel charts, which offer a static representation of articulation and fail to capture the dynamic nature of speech production. This limitation is particularly evident in representing the tongue position inside the mouth, as indicated by the frontness feature. Furthermore, recent research has shown that perception-based instruction is more effective than production-based methods such as corrective feedback [2]. The current study implemented a perception-based approach enhanced by rtMRI for teaching EFL vowels, a method that is supported both theoretically and empirically, as discussed in the following section.

1.1. Theoretical Framework

Perception-based pronunciation instruction has a theoretical basis in models of speech learning [3–5]. These models suggest that second language speech learning is perception-based. According to these models, perception is not only necessary for language comprehension but is also a precursor to language production. Similarly, empirical research suggests that perception-based instruction is effective [6]. Importantly, it can be more effective than production-based instruction [7]. In addition, stronger evidence for the effectiveness of perception-based pronunciation instruction comes from a meta-analytic review of a quarter of a century of research by Sakai and Moorman [8], which suggests that perception and production are connected. Therefore, if perception and production are closely related, then it stands that improving vowel perception is crucial for EFL learners, not only for comprehension but also to achieve intelligibility.

Similarly, the use of visual cues in second language speech learning is theoretically motivated. Liberman's Motor Theory of Speech Perception (e.g., Liberman & Mattingly [9]) posits that the perception of speech is closely linked to the motor processes involved in speech production. This theory suggests that listeners utilize their knowledge of how speech sounds are produced to understand what they hear. This theory has been implied by some models of speech perception and has been empirically supported (for a review see, Schwartz et al. [10] and Galantucci et al. [11]).

The Motor Theory framework is largely implied by the Multi-Cognitive Approach to teaching pronunciation [12]. This approach is grounded in the idea that effective pronunciation teaching must engage multiple cognitive and sensory channels to facilitate better learning outcomes. Odisho's approach is particularly relevant for learners who struggle with perceiving and producing sounds that are not present in their native language, as it emphasizes the integration of various sensory modalities to enhance the learning process. This approach contrasts with traditional methods that rely heavily on auditory input alone.

Visual input is a crucial element of the Multi-Cognitive Approach. Odisho underscores the significance of incorporating visual aids in learning speech sounds, such as articulatory diagrams, lip-reading, and advanced technologies like ultrasound and MRI. Visual feedback plays a vital role in helping learners comprehend the physical production of sounds by illustrating the movements of the articulators (tongue, lips, jaw, etc.). This is especially beneficial for sounds that pose challenges in perception or production solely through auditory input.

In support of the Multi-Cognitive Approach, empirical evidence suggests that L2 visual and audio-visual cues can improve both the perception and production of speech sounds[13–16]. In the current study, we aim to investigate the effect of using rtMRI videos in enhancing Arabic speakers' perception of EFL vowel sounds. No research exists on the direct application of rtMRI to vowel perception in EFL. This study builds upon existing literature to explore novel avenues for integrating technology into language education.

1.2. Phonological Differences between Arabic and English

Analyzing the sound system of L1 Arabic as compared to English requires an understanding of the phonological and phonetic foundations of this sound system. Unlike other EFL learners, where the L1 is typically the main factor in the perception or pronunciation difficulty of EFL sounds, the situation in Arabic is more complex. "Arabic" is technically an umbrella term under which different regional dialects may be encompassed. The discussion of the phonetic and phonological differences between these dialects is not within the scope of the current paper. Instead, two main Saudi Arabic dialects, namely Hijazi Arabic (HA) and Najdi Arabic (NA) dialects, are considered. The selection of these two dialects is crucial for two reasons: 1) Both dialects have been frequently investigated (e.g., Alghmaiz[17]; Alfaifi[18]; Alwazna[19]; Ammar & Alhumaid[20]), providing empirical linguistic data for comparison rather than relying on researchers' intuitions about sound inventories; and 2) HA and NA dialects are spoken by the participants in the current study, making it essential to understand the sound inventories of these two dialects.

However, the sound system of Arabic may not be influenced solely by the sound system of the native dialect (i.e., NA or HA) but, as noted by Alwazna[19], is also partly influenced by the sound system of Modern Standard Arabic (MSA). This is a result of the diglossic situation of the Arabic language, where Arabic speakers use their native dialect and MSA simultaneously (See Boudelaa and Marslen-Wilson[21] for a discussion of how the two varieties may be processed and represented). Therefore, the aim of this section is to compare and contrast the vowel systems of HA, NA, and MSA on one hand, and the vowel system of English on the other, to better understand the roots of pronunciation difficulties in EFL.

The pronunciation difficulties of EFL by Arabic speakers at the vocalic level are often attributed to the fact that Arabic has a smaller vowel inventory compared to English. **Table 1** highlights the differences and similarities in the vocalic systems of the two languages.

Table 1. A comparison of the vowel systems of English and some Arabic dialects.

Language	Short Vowels	Long Vowels	Diphthongs
MSA, NA, and HA[22–24]	/i, a, u/	/i:, a:, u:/	/ay/ and /aw/ in MSA. These two diphthongs, however, undergo monophthongization in both HA and NA dialects, changing their pronunciation to /e:/ and /o:/, respectively.
British English[1]	/ɪ, e, æ, ʌ, ɒ, ʊ, ə/	/i:, ɑ:, ɔ:, u:, ɜ:/	/eɪ, aɪ, ɔɪ, əʊ, aʊ, eə, ɪə, ʊə/

Furthermore, it was found that phonetically, English word stress is marked differently by Arabic-speaking learners of EFL compared to English native speakers[25]. In other words, Arabic speakers do not use vowel reduction to mark unstressed syllables in English as Native English speakers do.

1.3. Effects of Phonological Differences on Arabic-Speaking Learners of EFL

Pronunciation challenges encountered by Arabic-speaking EFL learners have been empirically documented, primarily focusing on the segmental level. Research has demonstrated that Arabic-speaking EFL learners struggle with certain consonant sounds that do not exist in their native language, such as /p/ and /v/[26, 27]. These challenges persist even among advanced learners when speaking rapidly, evident in tasks like tongue twisters where /p/ may be pronounced as /b/ and /v/ as /f/[28]. Additionally, Ammar and Alhumaid identified /p/, /ɹ/, /ʒ/, and /ŋ/ as the most problematic consonants for Arabic speakers, often substituted with /b/, the rolled /r/, /dʒ/, and /ng/ respectively[20, 29].

Recently, Rehman et al. utilized a database of read sentences to identify the segmental errors in the English pronunciation of four Arabic speakers[30]. A Saudi Arabic-speaking participant who lived in the United States for six years made 30% of phoneme substitution errors, 45% of deletion errors, and 18% of insertion errors. The latter included insertion errors involving final /ŋ/ pronounced as either /ŋg/ or /ŋk/ and vowel insertions in consonant clusters. Interestingly, the problematic English consonants, particularly /p/ and /v/, seem to be unaffected by the learning environment as they persist in both EFL settings[31] and ESL ones[28, 30].

Similar difficulties in vowel perception and production have also been observed. Studies on Arabic-speaking learners of English have highlighted the specific challenges they face in distinguishing between certain English vowel pairs, such as /ɪ/ (as in "pit") and /e/ (as in "pet"). For example, Altaha has shown that Arabic speakers substitute the English vowel /e/ with the vowel /ɪ/, pronouncing a word like *pen* /pen/ as /pin/[31]. Additionally, AlJasser has observed that Saudi English speakers tend to substitute the English diphthong /əʊ/ with the long vowel /u:/, making words like *home* and *whom* homophones[29].

Importantly, Arabic-speaking EFL learners appear to utilize a repair strategy when encountering consonant clusters that are not present in their native language, known as epenthesis or vowel insertion. For instance, in Rehman et al.'s study[30], a saudi participant accounted for 18% of all insertion errors among the four subjects, including instances of vowel insertion in consonant clusters (e.g., pronouncing 'spring' as /sɪpring/). Likewise, Alezetes discovered that Arabic-speaking English learners in her research, predominantly native speakers of Najdi Arabic, employed epenthesis to break up word-medial consonantal clusters, as seen in the word 'children'[32]. In a separate investigation, Arabic speakers, also native speakers of Najdi Arabic, faced challenges with English consonant clusters and consequently inserted epenthetic vowels both initially and finally in words[20].

Finally, Arabic-speaking English learners exhibit pronunciation errors at the suprasegmental level. Specifically, these learners tend to misplace English word stress due to the differences in suprasegmental systems between Arabic and English as discussed earlier. For example, Abker demonstrated that Arabic-speaking English learners often misplace word stress in both verbs and nouns[33]. Additionally, as highlighted by Almbark et al.[25], vowel reduction is not a common occurrence in the speech of Arabic-speaking English learners when producing unstressed syllables in English. Consequently, pronunciations such as [hi: went tu: sku:l] are frequently observed.

The studies discussed above highlight the challenges that Arabic-speaking learners of English face in the perception and production of English, particularly due to the differences between the vowel systems of Arabic and English. RtMRI offers a novel approach to addressing these challenges by providing learners with visual representations of articulatory movements, which can help them develop more accurate mental representations of English vowel sounds.

1.4. The Use of Technology-Enhanced Visual Feedback in the Teaching of L2 Perception and Production

The use of technology-enhanced visual feedback in the teaching of L2 phonology has gained increasing attention in recent years. Among the most promising technologies is ultrasound imaging which allows learners to visualize the articulatory processes involved in speech production[34–37]. Particularly, visualization of the tongue is provided. This section reviews key studies that have investigated the use of ultrasound in language teaching, with a focus on their effectiveness in improving L2 learners' perception and production of difficult phonetic contrasts, such as vowels and consonants.

D'Apolito et al. examined the impact of perceptual and ultrasound articulatory training on the pronunciation of

English vowels among Italian learners, specifically focusing on the challenging L2 vowel sounds /ʌ/, and /ɑ/[38]. The participants were divided into two experimental groups and one control group, each comprising only 3 participants. One experimental group underwent perceptual training that emphasized auditory discrimination tasks to improve their ability to differentiate these vowel sounds, while the other group received ultrasound articulatory training, which offered visual feedback on tongue placement during vowel production. The findings revealed that both training approaches resulted in notable enhancements in vowel accuracy, with ultrasound training showing particularly significant effects by enabling learners to visualize and adjust their articulatory positions, thereby aiding in improved pronunciation of the targeted vowels.

Similarly, Antolík et al. investigated the effectiveness of ultrasound on L2 pronunciation training of the French vowels /y/ and /u/. Four adult Japanese-speaking French L2 learners received three 45-minute sessions incorporating ultrasound pronunciation training[39]. Two other participants were used as controls. The analysis of the articulatory data indicated that three of the experimental participants showed significant improvement in their production of the French vowels, including the ability to distinguish between them and to contrast these sounds with the Japanese [ɯ]. These findings suggest that ultrasound may serve as an effective tool for enhancing pronunciation learning in second language acquisition.

More recently, Bryfonski investigated the comparative effects of visual and oral corrective feedback (CF) on the pronunciation development of second language (L2) learners[40]. Twenty-one Japanese learners of English were assigned to two groups: one receiving oral CF and the other receiving visual feedback through ultrasound. Participants were provided feedback on the English sounds /l/ and /r/ following induced communication breakdowns during task-based interactions. The effectiveness of the feedback was measured using accuracy ratings of /l/ and /r/ productions and comprehensibility ratings by naïve raters, both before and after the interactions. The results revealed that the ultrasound group showed significant gains in pronunciation accuracy during a story reading task compared to the oral CF group, although no differences were noted in a word list task. Additionally, participants in the ultrasound feedback group reported that the training was engaging and beneficial for their production of /r/ and /l/, highlighting the motivating effects of ultrasound visual feedback in pronunciation training.

Despite its potential benefits, the application of ultrasound technology may pose challenges for students, particularly in larger groups or for independent learners, due to difficulties in interpretation and the requirement for specialized equipment and expertise. To address these limitations, the development of ultrasound overlay videos has been proposed. Overlay videos are created by incorporating externally profiled views of the head with ultrasound images of speech tongue movements. A study by Alshehri comparing the production of the phonemes /p/ and /b/ among Arabic-speaking EFL learners who received ultrasound overlay video training versus those who did not found no significant differences in perception or pronunciation before, immediately after, and 11 days post-training[41]. These findings suggest that the use of ultrasound overlay videos may not effectively enhance the ability of young learners to perceive and pronounce these segments, nor to retain the benefits of the instruction over time. This raises questions about the practical application and efficacy of ultrasound technology in language teaching contexts.

The lack of conclusive evidence of the effectiveness of using ultrasound in speech learning is also supported by a systematic review of 28 papers investigating the effectiveness of ultrasound visual biofeedback in the treatment of speech sound disorders[42]. Their review showed that 79.3% of the studies represented lower levels of evidence. According to Sugden et al., lower strength designs and small sample sizes may have contributed to these findings. Arguably, other reasons may have contributed to the lack of significant effects of ultrasound treatments in some of these studies. First, as Antolík has noted, in some studies learners were learning new speech sounds that were not part of an L2 phoneme inventory[43]. So, their results cannot be generalized to L2 contexts. Second, others have focused on sounds that may be more difficult to teach using visual feedback alone, such as consonants with less obvious tongue involvement (e.g., Alshehri[41]). Additionally, individual differences in learners' ability to interpret the ultrasound videos or images may have contributed to the lack of significant effects in some cases, especially since most studies have used a small number of participants.

2. The Current Study

Taken together, the discussion above underscores the need for well-designed, large-sample studies to investigate the effectiveness of using technology-based visual cues in learning speech sounds. Additionally, the focus on the direct impact of technology-based visual cues in speech learning has overshadowed research on its potential effects on perception-based production. Moreover, learners' challenges in interpreting ultrasound videos or images may hinder their effectiveness. While ultrasound offers real-time feedback on tongue position, it lacks the anatomical detail of rtMRI. Critics, such as Gick et al.[34]), have highlighted the potential difficulty of interpreting ultrasound images, particularly for learners unfamiliar with the technology. The two-dimensional nature of ultrasound images may also restrict the comprehension of three-dimensional articulatory movements.

RtMRI can provide a significant improvement over ultrasound in terms of clarity. It offers high-resolution images of the entire vocal tract, including the tongue, velum, and pharynx, which are essential for understanding the articulatory movements involved in producing specific speech sounds. This comprehensive view enables educators to demonstrate not only tongue position but also the spatial relationships among various articulators. This dynamic visualization is particularly advantageous for teaching complex sounds that necessitate precise movement coordination. **Figure 1** compare still frames from ultrasound and MRI of the same person producing a /t/ sound on different occasions.

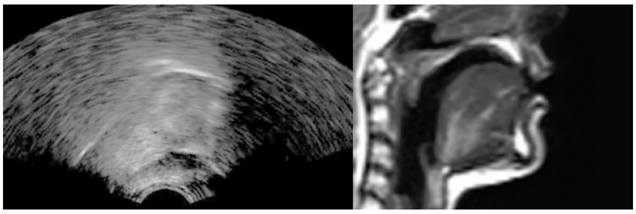

Figure 1. Comparison of Ultrasound (on the left) and MRI (on the right) still frames of the pronunciation of /t/. Adapted from "seeingspeech", available from: https://www.seeingspeech.ac.uk/mri-and-uti/.

Obviously, rtMRI may provide a more comprehensive and visually informative tool for teaching speech sounds, especially in a classroom setting where detailed anatomical understanding is beneficial. The drawback, however, is that MRI equipment is expensive. Fortunately, some online MRI resources have been developed recently. These include rtMRI videos of human speakers' pronunciations of International Phonetic Alphabet (IPA) sounds. A description of the tool used in the current study is provided in the methodology section. Against this background, the present study aims to:

- Examine the impact of integrating rtMRI training with EFL learners to enhance their comprehension of the articulatory movements linked to English vowels, enabling them to establish stronger connections between auditory cues and vowel features.
- Investigate the effectiveness of using rtMRI videos in enhancing Arabic speakers' perception of EFL vowel sounds.

Specifically, the study attempts to answer the following research questions:

1- Does training using rtMRI, as compared to traditional perception-based instruction, help Arabic-speaking EFL learners associate auditory cues with EFL vowel features?

2- Does training using rtMRI, as compared to traditional perception-based instruction, help Arabic-speaking EFL learners better perceive EFL vowel sounds?

3. Method

3.1. Participants

A total of forty male undergraduate students, all native speakers of Najdi Arabic, voluntarily participated in the study. They were randomly divided into two sections, with 20 students in each section, as part of the English pronunciation course in the English Language Program at a university in Saudi Arabia. The participants' ages ranged from 18 to 21 years, with a mean age of 19.3. None of the participants reported any hearing, speech, or vision problems.

3.2. Procedure

3.2.1. Pre-Tests

Two tests were utilized in the study. The first test was the vowel-feature association test (VFAT), focusing specifically on the frontness feature to assess participants' visual representation of tongue shape during vowel perception. The VFAT aimed to determine if instruction with rtMRI improved participants' ability to visually represent tongue movements during vowel perception, shedding light on whether any perceived lack of effect on vowel perception was due to a failure in visual representation rather than a failure in transferring that representation to vowel perception. Participants were required to listen to various English vowels in isolation and select the corresponding tongue position from *front, central,* and *back* options for each vowel. The VFAT stimuli included eleven RP vowels, comprising of four front vowels (/ɪ, e, æ, iː/), two central vowels (/ʌ, ɜː/), and five back vowels (/ɒ, ʊ, ɑː, ɔː, uː/). These vowels were recorded by a female native speaker of Standard Southern British English (SSBE) using a high-quality microphone and saved onto a computer disk at a sampling rate of 44.1 kHz.

The second assessment was a vowel perception test (VPT). In this test, participants listened to the same vowel sounds used in the previous test, but within the context of /h/-V-/d/ words (e.g., had, heed). Their task was to identify each word presented aurally from a written list of options, which included one correct answer (e.g., had, heed, hard, heard). To ensure that our EFL participants were familiar with the words and to avoid influencing the results, we included a commonly rhyming word in each option, as suggested by Evans and Alshangiti[44] (e.g., hid as in kid). In the VPT, the same 11 vowels from the VFAT were used within the context of CVC (i.e., /h/-V-/d/) words (e.g., had, heed). The test consisted of eleven /h/-V-/d/ words, embedded in the sentence "I say a ____." These sentences were recorded by the same female native speaker of Standard Southern British English (SSBE) using the same procedure as in the recording of the VFAT stimuli.

3.2.2. Intervention

Experimental group: Participants in this group received vowel perception training for three weekly 30-minute sessions using rtMRI videos. This included watching rtMRI visualizations videos of tongue, lips, and jaw movements while listening to target vowels. This is in addition to traditional perception-based training as in with the control group. In each session, a group of English vowels were practiced. At the start of each session, the teacher would display the rtMRI video on the projector screen with the sound on. The second phase involved demonstrating rtMRI videos with the sounds muted, and students' task involved guessing which vowel is being pronounced based on the muted rtMRI video.

The website used for MRI training was created by the speech production and articulation knowledge group (SPAN) at the University of Southern California. One of the rtMRI resources available on the website which can be accessed at https://sail.usc.edu/span/rtmri_ipa/) makes available "..rtMRI data corresponding to a large sub-set of the sounds of the world's languages as encoded in the International Phonetic Alphabet, with supplementary English words and phonetically-balanced texts, produced by four prominent phoneticians, using the latest rtMRI technology"[45].

Control group: Participants received traditional perception training using vowel charts, and explicit verbal descriptions of the target articulatory movements.

3.2.3. Post-Tests

All participants underwent post-testing using the same VFAT and VPT assessments. The post-test was conducted immediately after the final training session, with procedures identical to those of the pre-tests.

4. Results

4.1. Hypothesis 1

EFL learners who receive training using rtMRI videos will demonstrate a significantly better ability to associate auditory cues with EFL vowel features compared to those who receive traditional perception-based instruction.

To test this hypothesis, we utilized an independent-samples t-test to compare the performance of the experimental group (rtMRI training) and the control group (traditional perception-based instruction) in associating auditory cues with EFL vowel features. The t-test was employed to determine if there exists a statistically significant variance in the mean scores between the two groups.

The descriptive statistics in **Table 2** suggest that the experimental group made substantial improvements in the posttest (Mean = 8.05), whereas the control group showed only marginal improvement (Mean = 4.85). Additionally, both groups exhibited acceptable levels of skewness and kurtosis, suggesting that the data is roughly normally distributed.

As shown in **Table 3**, the results of the Kolmogorov-Smirnov and Shapiro-Wilk tests indicate that the pretest and posttest scores for both the experimental and control groups are normally distributed. None of the p-values are below the

Table 2. Descriptive statistics for normality checks on VFAT pretest and posttest scores.

Test	Group	Mean	Std. Deviation	Skewness	Kurtosis
Pretest	Experimental	4.20	1.704	0.079	−1.166
	Control	4.10	1.586	0.346	−0.667
Posttes	Experimental	8.05	1.504	0.526	−0.313
	Control	4.85	1.755	0.513	−0.013

0.05 threshold, suggesting that the assumption of normality is not violated. This supports the use of parametric tests such as the independent-samples t-test for comparing the two groups.

Table 3. Tests of normality for VFAT pretest and posttest scores for both groups.

Test	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
pretest	Experimental	0.152	20	0.200*	0.911	20	0.065
	Control	0.175	20	0.109	0.922	20	0.108
posttest	Experimental	0.163	20	0.170	0.922	20	0.109
	Control	0.186	20	0.068	0.941	20	0.252

Table 4 presents both descriptive statistics and results from the independent samples t-test to compare the pretest scores between the experimental and control groups. Here are the key points:

- **Front Vowels:** Both groups have similar means (1.30 for the experimental group and 1.40 for the control group). The t-test result (t = −0.362, p = 0.719) indicates no significant difference between them. This is supported by the 95% confidence interval (−0.659 to 0.459), which includes zero, suggesting that the groups are equivalent.
- **Central Vowels:** The mean scores are similar (0.90 for the experimental group and 1.05 for the control group), and the t-test result (t = −0.761, p = 0.451) confirms no significant difference. The confidence interval (−0.549 to 0.249) further supports the statistical similarity between the groups.
- **Back Vowels:** Although the experimental group has a slightly higher mean (2.00) compared to the control group (1.65), the t-test result (t = 0.892, p = 0.378) indicates no significant difference. The confidence interval (−0.445 to 1.145) further supports this finding.
- VFAT **Pretest Total**: The total VFAT pretest scores are nearly identical (4.20 for the experimental group and 4.10 for the control group). The t-test result (t = 0.192, p = 0.849) shows no significant difference between the groups, with a confidence interval (−0.954 to 1.154), confirming that the groups were well-matched at the start of the study.

Table 4. Independent samples t-test results for VFAT pretest equivalence.

Variable	Group	N	Mean	Std. Deviation	Levene's Test (F)	Levene's Sig.	t-Value	df	Sig. (2-Tailed)
Front Vowels	Experimental	20	1.30	0.733	1.390	0.246	−0.362	38	0.719
	Control	20	1.40	0.995					
Central Vowels	Experimental	20	0.90	0.641	0.238	0.628	−0.761	38	0.451
	Control	20	1.05	0.605					
Back Vowels	Experimental	20	2.00	1.376	0.975	0.330	0.892	38	0.378
	Control	20	1.65	1.089					
Pretest Total	Experimental	20	4.20	1.704	0.423	0.519	0.192	38	0.849
	Control	20	4.10	1.586					

In conclusion, the t-test results for all variables indicate that there are no statistically significant differences between the experimental and control groups prior to the intervention. This provides a solid baseline for assessing post-intervention outcomes.

As shown in **Table 5**, the results of the Independent Samples T-Test after the intervention reveal statistically significant differences between the experimental and control groups across all measured variables, with corresponding effect sizes (Eta Squared) providing insights into the strength of these differences.

- **Front Vowels:** The experimental group (Mean = 2.95) significantly outperformed the control group (Mean = 1.45), with a t-value of 4.880 (p = 0.000). The Eta Squared value ($\eta^2 = 0.392$) indicates a large effect size, demonstrating that the rtMRI training had a substantial impact on the VFAT scores related to front vowels. The confidence interval (0.878 to 2.122) further strengthens the evidence of this significant effect.
- **Central Vowels:** The experimental group (Mean = 1.55) significantly outperformed the control group (Mean = 0.85), as indicated by a t-value of 3.090 (p = 0.004). The Eta Squared value ($\eta^2 = 0.197$) signifies a moderate effect size, highlighting the efficacy of rtMRI training on the VFAT scores related to central vowels. The confidence interval (0.241 to 1.159) supports these results.
- **Back Vowels**: The experimental group (Mean = 3.55) significantly outperformed the control group (Mean = 2.55), with a t-value of 2.562 (p = 0.015). The Eta Squared value ($\eta^2 = 0.144$) indicates a moderate effect size, showing that the training had a noticeable impact on the VFAT scores related to back vowels. The confidence interval (0.210 to 1.790) further reinforces this finding.
- **Post-test Total**: The overall scores of the experimental group (Mean = 8.05) were significantly higher, as depicted in **Figure 2**, than those of the control group (Mean = 4.85), with a t-value of 6.192 (p = 0.000). The large effect size ($\eta^2 = 0.568$) highlights the significant impact of the rtMRI training on the overall vowel-feature association ability. The confidence interval (2.154 to 4.246) further supports this strong positive effect.

Table 5. Independent samples t-test results for VFAT posttest differences.

Variable	Group	N	Mean	Std. Deviation	Levene's Test (F)	Levene's Sig.	t-Value	df	Sig. (2-Tailed)	Eta Squared (η^2)
Front Vowels	Experimental Control	20 20	2.95 1.45	0.887 1.050	0.857	0.360	4.880	38	0.000	0.392
Central Vowels	Experimental Control	20 20	1.55 0.85	0.605 0.813	1.685	0.202	3.090	38	0.004	0.197
Back Vowels	Experimental Control	20 20	3.55 2.55	0.999 1.432	3.323	0.076	2.562	38	0.015	0.144
Posttest Total	Experimental Control	20 20	8.05 4.85	1.504 1.755	0.955	0.335	6.192	38	0.000	0.568

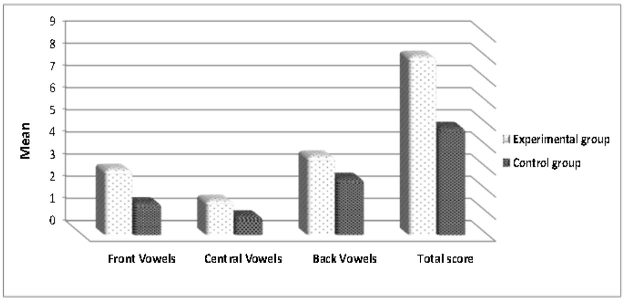

Figure 2. Comparison of posttest scores for VFAT between the experimental group and the control group.

In summary, these findings demonstrate that the rtMRI training significantly enhanced the ability to associate vowels with their features across all frontness categories compared to traditional instruction methods, as supported by the significant t-test results and the substantial effect sizes (Eta Squared).

4.2. Hypothesis 2

EFL learners who receive training using rtMRI videos will demonstrate significantly better perception of EFL vowel sounds compared to those who receive traditional perception-based instruction.

To test this hypothesis, we utilized an independent-samples t-test to compare the performance of the experimental group and the control group on their ability to perceive

EFL vowel sounds in the VPT.

Table 6 displays the means and standard deviations for VPT pretest and posttest scores of both groups. Posttest scores indicate improved performance in both groups, with the experimental group achieving a higher mean (8.70) compared to the control group (8.20). Skewness values suggest a slight leftward asymmetry in both pretest and posttest distributions, while kurtosis values indicate varying distribution shapes. Overall, these results suggest both groups improved, with the experimental group showing a slightly bigger improvement. We will now analyze whether this improvement was statistically significant.

Table 6. Descriptive statistics for normality checks on VPT pretest and posttest scores.

Test	Group	Mean	Std. Deviation	Skewness	Kurtosis
Pretest	Experimental	6.75	2.531	−0.813	1.267
	Control	7.35	2.907	−0.658	−0.347
Posttes	Experimental	8.70	1.490	−0.167	−0.410
	Control	8.20	1.989	−0.930	1.174

Table 7 presents the results of normality tests for both pretest and posttest scores using the Kolmogorov-Smirnov and Shapiro-Wilk tests. For the pretest, the experimental group shows a K-S statistic of 0.189 with a p-value of 0.059, indicating a marginal departure from normality, while the control group's K-S statistic of 0.152 and p-value of 0.200 suggests normality. For the posttest, the experimental group also displays a K-S statistic of 0.180 and a p-value of 0.090, further indicating a potential deviation from normality. However, the control group's posttest results (K-S = 0.160, p = 0.193) suggest that the distribution remains approximately normal. Overall, these results imply that while some deviations from normality exist, the control group demonstrates more consistent normality in both test phases.

Table 7. Tests of normality for VPT pretest and posttest scores.

Test	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
pretest	Experimental	0.189	20	0.059	0.933	20	0.178
	Control	0.152	20	0.200*	0.932	20	0.171
posttest	Experimental	0.180	20	0.090	0.931	20	0.159
	Control	0.160	20	0.193	0.931	20	0.161

Table 8 presents both descriptive statistics and results from the independent samples t-test to compare the pretest scores between the experimental and control groups. Here are the key points:

Front Vowels: The experimental group had a mean score of 2.65, while the control group scored 3.20. The t-test result (t = −1.621, p = 0.113) indicates no significant difference between the groups. The confidence interval (−1.237 to 0.137) includes zero, supporting that the groups were equivalent in their perception of front vowels prior to the intervention.

Central Vowels: The experimental group had a mean score of 1.50, while the control group scored 1.45. The t-test result (t = 0.230, p = 0.819) indicates no significant difference between the groups. The confidence interval (−0.390 to 0.490) supports that both groups were statistically similar in their perception of central vowels prior to the intervention.

Back Vowels: The experimental group had a mean score of 2.60, slightly lower than the control group's mean of 2.70. Nevertheless, the t-test result (t = −0.209, p = 0.836) shows no significant difference between the groups. The confidence interval (−1.069 to 0.869) including zero indicates that the groups were equivalent in their perception of back vowels.

Pretest Total: The overall pretest scores for both groups were similar, with the experimental group scoring 6.75 and the control group scoring 7.35. The t-test result (t = −0.696, p = 0.491) indicates no significant difference

Table 8. Independent samples t-test results for VPT pretest equivalence.

Variable	Group	N	Mean	Std. Deviation	Levene's Test (F)	Levene's Sig.	t-Value	df	Sig. (2-Tailed)
Front Vowels	Experimental	20	2.65	1.089	0.040	0.842	−1.621	38	0.113
	Control	20	3.20	1.056					
Central Vowels	Experimental	20	1.50	0.607	1.442	0.237	0.230	38	0.819
	Control	20	1.45	0.759					
Back Vowels	Experimental	20	2.60	1.273	2.692	0.109	−0.209	38	0.836
	Control	20	2.70	1.720					
Pretest Total	Experimental	20	6.75	2.531	0.396	0.533	−0.696	38	0.491
	Control	20	7.35	2.907					

between the groups, and the confidence interval (−2.345 to 1.145) confirms that the groups were well-matched at the beginning of the study.

These results indicate that there were no significant pre-existing differences between the experimental and control groups, ensuring a fair comparison for the subsequent intervention effects.

As shown in **Table 9**, the results of the independent samples t-test for the second research question indicate no significant differences between the experimental and control groups on the posttest scores. Specifically:

- **Front Vowels**: The t-value of 0.167 and p-value of 0.868 reveal no statistically significant difference between the experimental and control groups. The confidence interval (−0.555 to 0.655) encompasses zero, suggesting that both groups performed similarly in perceiving front vowels.
- **Central Vowels**: A t-value of 0.309 and a p-value of 0.759 reveal no significant difference between the groups on central vowel scores. The confidence interval (−0.278 to 0.378) further supports this, confirming the lack of a meaningful distinction between the two groups.
- **Back Vowels**: The t-value of 0.892 and p-value of 0.378 indicate no significant difference between the groups on back vowel scores. The confidence interval (−0.445 to 1.145) encompasses zero, suggesting that the experimental and control groups performed similarly.
- **Posttest Total**: The overall posttest scores, as depicted in **Figure 3**, show no significant difference, as indicated by a t-value of 0.900 and p-value of 0.374. The confidence interval (−0.625 to 1.625) includes zero, suggesting that there was no substantial difference in the total scores between the two groups.

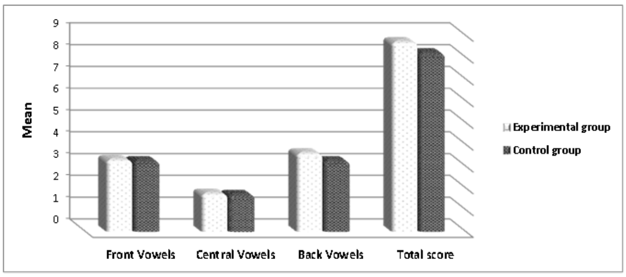

Figure 3. Comparison of posttest scores for VPT between experimental group and the control group.

These results suggest that rtMRI training did not significantly impact the learners' ability to perceive EFL vowels compared to traditional perception-based instruction, as both groups showed similar post-training performance across all vowel categories.

5. Discussion

The current research aimed to explore whether perceptual training using rtMRI can be superior to traditional perceptual training in developing the perception of EFL vowels. The results of this study provide valuable insights into the potential benefits of using rtMRI videos as a pedagogical tool for enhancing the perception of EFL vowel sounds among Arabic-speaking learners. The findings indicate that the experimental group, which received perception-based instruction with the aid of rtMRI videos, demonstrated significant improvement in associating vowel sounds with their articulatory features, while the control group, which received traditional instruction, did not show similar gains. This finding suggests that the use of rtMRI videos, which visually depict the internal articulatory movements during speech production, offers learners a more concrete understanding of how vowel sounds are produced. The visual representation of tongue positioning and other articulatory mechanisms likely

Table 9. Independent samples t-test results for VPT posttest differences.

Variable	Group	N	Mean	Std. Deviation	Levene's Test (F)	Levene's Sig.	t-Value	df	Sig. (2-Tailed)
Front Vowels	Experimental	20	3.30	0.865	1.090	0.303	0.167	38	0.868
	Control	20	3.25	1.020					
Central Vowels	Experimental	20	1.75	0.444	0.692	0.411	0.309	38	0.759
	Control	20	1.70	0.571					
Back Vowels	Experimental	20	3.60	1.095	1.582	0.216	0.892	38	0.378
	Control	20	3.25	1.372					
Posttest Total	Experimental	20	8.70	1.490	0.966	0.332	0.900	38	0.374
	Control	20	8.20	1.989					

helped learners form stronger mental associations between the auditory and visual aspects of vowel production.

This finding is in line with theoretical accounts that emphasize the importance of visualizing speech [9, 46]. Furthermore, it supports previous studies on visual feedback in language learning, which have demonstrated that learners benefit from visually observing how sounds are produced, especially when the sounds are challenging to perceive or produce due to differences in phonetic inventories between the learners' native language and the target language [38–40, 47]. This aligns with the findings of Cai et al. [48], which demonstrate that optimizing auditory input can significantly aid language learners in processing sounds. For Arabic-speaking learners, English vowel sounds, particularly those without direct equivalents in Arabic, can pose challenges in perception and production. The use of rtMRI videos likely offered a more intuitive way for learners to comprehend these unfamiliar sounds.

Interestingly, while the experimental group showed a significant improvement in the vowel-feature association test, the gains observed in the vowel perception test for the experimental group were minimal. This could suggest that while the rtMRI videos were effective in helping learners understand the articulatory features of vowel sounds, this understanding did not immediately translate into improved perceptual discrimination of the vowels themselves. One possible explanation for this is that the perception of vowel sounds, especially in a second language, is a complex process that involves not only articulatory awareness but also auditory discrimination skills that may take longer to develop. Previous research has shown that perceptual learning of non-native vowel contrasts is significantly influenced by L1 experience (e.g., Iverson et al. [49]) and therefore L2 sound learning can be a slow process, often requiring extensive exposure and practice. It is possible that the three 30-minute sessions in this study were not sufficient for learners to fully internalize the auditory distinctions between the vowel sounds, even though they had gained a better understanding of their articulatory features.

The findings of this study highlight the potential role of articulatory awareness in language learning, especially in the context of vowel perception. While traditional perception-based instruction typically emphasizes auditory discrimination alone, the incorporation of rtMRI videos introduces a novel dimension by elucidating the articulatory processes involved in vowel production. This could be particularly advantageous for learners whose native language phonetic system diverges significantly from that of the target language, as is the case with Arabic and English.

The significant improvement in vowel-feature association observed in the experimental group suggests that articulatory awareness can be a valuable complement to auditory training. By understanding how vowel sounds are produced, learners may be better equipped to produce these sounds themselves, even if their perceptual discrimination abilities are still developing. This is consistent with the findings of studies that have shown a positive relationship between perceiving speech and perceiving gestures (for a review, see Galantucci et al. [11]).

5.1. Implications for EFL Instruction

The results of this study have important implications for EFL instruction, particularly in the area of pronunciation and phonetic training. The use of rtMRI videos represents a novel approach that could be integrated into pronunciation curricula to enhance learners' understanding of vowel production. Given the significant gains observed in the vowel-feature association test, incorporating visual tools like rtMRI videos could help learners develop a more holistic understanding of vowel sounds, combining both auditory and articulatory information.

However, the lack of significant improvement in vowel perception suggests that rtMRI videos should be used as a complement to, rather than a replacement for, traditional auditory training. EFL instructors may need to provide learners with additional opportunities for auditory discrimination practice, possibly through extended exposure to the target vowel sounds in various phonetic contexts. Combining articulatory awareness with auditory training could lead to more robust gains in both perception and production.

5.2. Limitations and Future Research

While the findings of this study are promising, there are several limitations that should be acknowledged. First, the duration of the intervention was relatively short, with only three 30-minute sessions. It is possible that a longer intervention period would have led to greater gains in both vowel-feature association and vowel perception. Future studies could explore the effects of more extended training with rtMRI videos to determine whether longer exposure leads to more significant improvements in vowel perception.

Moreover, while this study focused on vowel perception, future research could investigate the effects of rtMRI videos on other aspects of pronunciation, such as consonant perception and production, or even prosodic features like intonation and stress. Expanding the scope of research to include different phonetic features could provide a more comprehensive understanding of the potential benefits of rtMRI videos in language learning.

6. Conclusions

In conclusion, this study provides evidence that rtMRI videos can be an effective tool for enhancing learners' understanding of the articulatory features of English vowel sounds. While the experimental group showed significant gains in associating vowel sounds with their articulatory features, no significant improvement was observed in vowel perception. These findings suggest that rtMRI videos can play a valuable role in developing articulatory awareness, but they should be used in conjunction with traditional auditory training to achieve more comprehensive gains in vowel perception. Future research should continue to explore the potential of rtMRI videos and other visual tools in language learning, with a focus on optimizing their use for different aspects of pronunciation and phonetic training. Further research is needed to explore how these technologies can be optimized for different learner populations.

Funding

The Researcher would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024).

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Qassim University (24-73-08/Jan18-2024).

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data that support the findings of this study are available from the author upon reasonable request.

Acknowledgments

The Researcher would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024).

Conflicts of Interest

The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] Roach, P., 2009. English Phonetics and Phonology, 4th ed. Cambridge University Press: Cambridge, UK. pp. 1–231.

[2] Lee, B., Plonsky, L., Saito, K., 2019. The Effects of Perception- Vs. Production-Based Pronunciation Instruction. System. 88, 1–13. DOI: https://doi.org/10.1016/j.system.2019.102185

[3] Best, C.T., 1995. A Direct Realist View of Cross-Language Speech Perception. In: Strange, W. (ed.). Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press: Timonium, US. pp. 171–204.

[4] Flege, J.E., 1995. Second Language Speech Learning: Theory, Findings and Problems. In: Strange, W. (ed.). Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press: Timonium, US. pp. 233–277.

[5] Best, C.T., Tyler, M.D., 2007. Nonnative and Second-Language Speech Perception: Commonalities and Complementarities. In: Munro, M.J., Bohn, O. (eds.). Language Experience in Second Language Speech Learning: In Honor of James Flege. John Benjamins: Amsterdam, Netherlands. pp. 13–34. DOI: https://doi.org/10.1075/lllt.17.07bes

[6] Thomson, R., 2011. Computer Assisted Pronunciation Training: Targeting Second Language Vowel Perception Improves Pronunciation. CALICO Journal. 28(3), 744–765.

[7] Lee, B., Plonsky, L., Saito, K., 2020. The Effects of Perception vs. Production-based Pronunciation Instruction. System. 88, 102185. DOI: https://doi.org/10.1016/j.system.2019.102185

[8] Sakai, M., Moorman, C., 2018. Can Perception Training Improve the Production of Second Language Phonemes? A Meta-Analytic Review of 25 Years of Perception Training Research. Applied Psycholinguistics. 39(1), 187–224. DOI: https://doi.org/10.1017/S0142716417000418

[9] Liberman, A., Mattingly, I., 1985. The Motor Theory of Speech Perception Revised. Cognition. 21(1), 1–36. DOI: https://doi.org/10.1016/0010-0277(85)90021-6

[10] Schwartz, J.L., Sato, M., Fadiga, L., 2009. The Common Language of Speech Perception and Action: A Neurocognitive Perspective. Revue Francaise de Linguistique Appliquee. 13(2), 9–22. DOI: https://doi.org/10.3917/rfla.132.0009

[11] Galantucci, B., Fowler, C., Turkey, M., 2006. The Motor Theory of Speech Perception Reviewed. Psychonomic Bulletin and Review. 13(3), 361–377. DOI: https://doi.org/10.3758/PBR.15.2.453

[12] Odisho, E.Y., 2007. A Multisensory, Multicognitive Approach to Teaching Pronunciation. Linguística - Revista de Estudos Linguísticos Da Universidade Do Porto. 2, 3–28.

[13] Hazan, V., Sennema, A., Faulkner, A., et al., 2006. The Use of Visual Cues in the Perception of Non-Native Consonant Contrasts. The Journal of the Acoustical Society of America. 119(3), 1740–1751. DOI: https://doi.org/10.1121/1.2166611

[14] Hazan, V., Sennema, A., Iba, M., et al., 2005. Effect of Audiovisual Perceptual Training on the Perception and Production of Consonants by Japanese Learners of English. Speech Communication. 47(3), 360–378. DOI: https://doi.org/10.1016/j.specom.2005.04.007

[15] Hong, S., Wang, R., Zeng, B., 2023. Incongruent Visual Cues Affect the Perception of Mandarin Vowel But Not Tone. Frontiers in Psychology. 13, 1–10. DOI: https://doi.org/10.3389/fpsyg.2022.971979

[16] Mahmood, R.Q., 2024. The Impact of Visual Corrective Feedback on Pronunciation Accuracy in L2 Sound Production: Empirical Evidence. In Bang, T., Nguyen, C., Bui, H. (eds.), Exploring Contemporary English Language Education Practices. IGI Global: Pennsylvania, US. pp. 158–189. DOI: https://doi.org/10.4018/979-8-3693-3294-8.ch008

[17] Alghmaiz, B., 2013. Word-Initial Consonant Cluster Patterns in the Arabic Najdi Dialect [Master Thesis]. Carbondale, IL: Southern Illinois University Carbondale. pp. 1–61.

[18] Alfaifi, A., 2019. Syllabification of Coda Consonant Clusters in Najdi and Hijazi Arabic [Doctoral Dissertation]. Fairfax, VA: George Mason University. pp. 1–124.

[19] Alwazna, R., 2020. The Differences in Consonantal Pronunciation Between Formal English and Saudi Hijazi English and Their Implications for Oral Intralingual Translation. Journal of Psycholinguistic Research. 49(4), 571–582. DOI: https://doi.org/10.1007/s10936-020-09697-5

[20] Ammar, W., Alhumaid, A., 2009. A Phonetic Study of Najdi Arabic Interference in L2 Acquisition of English. Arab Journal for the Humanities. 27(108), 213–235.

[21] Boudelaa, S., Marslen-Wilson, W.D., 2013. Morphological Structure in the Arabic Mental Lexicon: Parallels Between Standard and Dialectal Arabic. Language and Cognitive Processes. 28(10), 1453–1473. DOI: https://doi.org/10.1080/01690965.2012.719629

[22] Ingham, B., 1994. Najdi Arabic. John Benjamins: Amsterdam, US. pp. 1–215.

[23] Ryding, K., 2005. A Reference Grammar of Modern Standard Arabic. Cambridge University Press: Cambridge, UK. pp. 1–708.

[24] Watson, J.C.E., 2007. The Phonology and Morphology of Arabic. Oxford University Press: Oxford, UK. pp. 1–324.

[25] Almbark, R., Bouchhioua, N., Hellmuth, S., 2014. Acquiring the Phonetics and Phonology of English Word Stress: Comparing Learners From Different L1 Backgrounds. Proceedings of the International Symposium on the Acquisition of Second Language Speech; (June 26 - June 28, 2014); (Tainan, Taiwan). pp. 19–35.

[26] Flege, J.E., Port, R., 1981. Cross-Language Phonetic Interference: Arabic to English. Language and Speech.

24(2), 125–146.

[27] Alotaibi, A., 2013. Pronunciation Problems in the Production of the Voiced Labiodental Fricative /v/ by Saudi Speakers of English. [Master Thesis], Carbondale, IL: Southern Illinois University.

[28] Aljasser, F., Jackson, K.T., Vitevitch, M.S., et al, 2018. The Influence of Phoneme Inventory on Elicited Speech Errors in Arabic Speakers of English. The Mental Lexicon. 13(1), 26–37. DOI: https://doi.org/10.1075/ml.17014.alj

[29] Aljasser, J., 1978. Interlingual Errors and the Phonological Difficulties That Saudi Arabian Learners of English Encounter [Level of Thesis]. City, State Abbreviation: University of Kansas. pp. 1–53.

[30] Rehman, I., Silpachai, A., Levis, J., et al., 2020. The English Pronunciation of Arabic Speakers: A Data-Driven Approach to Segmental Error Identification. Language Teaching Research. 26(6), 1–27. DOI: https://doi.org/10.1177/1362168820931888

[31] Altaha, F.M., 1995. Pronunciation Errors Made by Saudi University Students Learning English. ITL - International Journal of Applied Linguistics. 109(1), 110–123.

[32] Alezetes, E., 2007. A Markedness Approach to Epenthesis in Arabic Speakers' L2 English [Master Thesis]. Missoula, MT: The University of Montana. pp. 1–103.

[33] Abker, I., 2019. Pronunciation Problems of Word Stress Placement in English by Saudi Students at Albaha University, Saudi Arabia: A Case Study at Almandag. International Journal of Applied Linguistics and English Literature. 8(6), 103–113. DOI: https://doi.org/10.7575/aiac.ijalel.v.8n.6p.103

[34] Gick, B., Bernhardt, B., Bacsfalvi, P., et al., 2008. Ultrasound Imaging Applications in Second Language Acquisition. In: Edwards, J., Zampini, M. (eds.). Phonology and Second Language Acquisition, 36th ed. John Benjamins Publishing Company: Philadelphia, US. pp. 309–322. DOI: https://doi.org/10.1075/sibil.36.15gic

[35] Holt, E., 2018. The Use of MRI and Ultrasound Technology in Teaching Pronunciation. Proceedings of the 9th Annual Pronunciation in Second Language Learning and Teaching Conference; Date of Conference (August 31- September 2, 2017); University of Utah, USA (Salt lake, UT). pp. 252–262. Available from: https://apling.engl.iastate.edu/wp-content/uploads/sites/221/2018/09/PSLLT-Proceedings-9_9-14-18_2.pdf (cited 11 October 2024).

[36] Preston, J.L., Byun, T.M., Boyce, S.E., et al., 2017. Ultrasound Images of the Tongue: A Tutorial for Assessment and Remediation of Speech Sound Errors. Journal of Visualized Experiments. 119, 1–10.

[37] Wilson, I., 2014. Using Ultrasound for Teaching and Researching Articulation. Acoustical Science and Technology. 35(6), 285–289. DOI: https://doi.org/10.1250/ast.35.285

[38] D'Apolito, I., Sisinni, B., Grimaldi, M., et al., 2017. Perceptual and Ultrasound Articulatory Training Effects on English L2 Vowels Production by Italian Learners. International Journal of Cognitive and Language Sciences. 11(8), 2174–2181.

[39] Antolík, T., Pillot-Loiseau, C., Kamiyama, T., 2019. The Effectiveness of Real-Time Ultrasound Visual Feedback on Tongue Movements in L2 Pronunciation Training: Japanese Learners' Progress on the French Vowel Contrast /y/-/u/. Journal of Second Language Pronunciation. 5(1), 72–97.

[40] Bryfonski, L., 2023. Is Seeing Believing?: The Role of Ultrasound Tongue Imaging and Oral Corrective Feedback in L2 Pronunciation Development. Journal of Second Language Pronunciation. 9(4), 103–109. DOI: https://doi.org/10.1075/jslp.22051.bry

[41] Alshehri, F.H., 2022. Ultrasound Overlay Video Effectiveness in Teaching Pronunciation to Young EFL Learners. Theory and Practice in Language Studies. 12(8), 1545–1555.

[42] Sugden, E., Lloyd, S., Lam, J., et al., 2019. Systematic Review of Ultrasound Visual Biofeedback in Intervention for Speech Sound Disorders. International Journal of Language and Communication Disorders. 54(5), 705–728.

[43] Antolík, T., 2020. Ultrasound Tongue Imaging in Second Language Learning. Studiez Aplikované Lingvistiky. 11(1), 109–119.

[44] Bronwen, E., Alshangiti, W., 2018. The Perception and Production of British English Vowels and Consonants by Arabic Learners of English. Journal of Phonetics. 68, 15–31.

[45] Toutios, A., Lingala, S.G., Vaz, C., et al., 2016. Illustrating the Production of the International Phonetic Alphabet Sounds Using Fast Real-Time Magnetic Resonance Imaging. Proceedings of the 17th Annual Conference of the International Speech Communication Association, INTERSPEECH; Date of Conference (September 8- September 12, 2016); USA (San Francisco, Cal). pp. 2428–2432.

[46] Odisho, E.Y., 2007. A Multisensory, Multicognitive Approach to Teaching Pronunciation. Linguística - Revista de Estudos Linguísticos Da Universidade Do Porto. 2, 3–28.

[47] Hardison, D., 2003. Acquisition of Second-Language Speech: Effects of Visual Cues, Context, and Talker Variability. Applied Psycholinguistics. 24(4), 495–522.

[48] Cai, X., Lian, A., Puakpong, N., et al., 2021. Optimizing Auditory Input for Foreign Language Learners Through a Verbotonal-Based Dichotic Listening Approach. Asian-Pacific Journal of Second and Foreign Language Education. 6(14). DOI: https://doi.org/10.1186/s40862-021-00119-0

[49] Iverson, P., Kuhl, P., Akahane-Yamada R., et al., 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition. 87(1), B47–B57.